

Compensating asynchrony effects in the calculation of financial correlations

Michael C. Münnix*, Rudi Schäfer, Thomas Guhr

Fakultät für Physik, Universität Duisburg-Essen, Germany

Abstract

We present a method to compensate statistical errors in the calculation of correlations on asynchronous time series. The method is based on the assumption of an underlying time series. We set up a model and apply it to financial data to examine the decrease of calculated correlations towards smaller return intervals (Epps effect). We show that this statistical effect is a major cause of the Epps effect. Hence, we are able to quantify and to compensate it using only trading prices and trading times.

Keywords: Financial correlations, Epps effect, Market emergence, Covariance estimation, Asynchronous time series

1. Introduction

The decrease of calculated correlations in financial data towards smaller return (or “sampling”-) intervals has been of interest since Epps discovered this phenomenon in 1979 [1]. Ever since, this behavior was found in data of different stock exchanges [2, 3, 4, 5] and foreign exchange markets [6, 7].

Many economists as well as physicists addressed this phenomenon, since a precise calculation of correlations is of major importance for the estimation of financial risk [8, 9, 10]. While the physicists’ approach is often to construct a model which offers an explanation for this phenomenon, the standard economy approach is to work on estimators with the aim to suppress the Epps effect. Recently, Hayashi and Yoshida introduced a cumulative estimator [11], only involving returns whose time intervals are overlapping. This estimator has been supplemented with different adjustments, such as bias compensation and *lead-lag* treatment [5, 12, 13]. A very similar approach on a completely different topic is the “discrete correlation function” in astrophysics which was introduced in 1988 by Edelson and Krolik [14]. Other approaches to estimate correlation coefficients involve *Previous-Tick-Estimators* [15, 16] or realized kernel functions [17].

An extensive study of microscopic causes leading to the Epps effect has been performed by Renò [18], while another work by Tóth et. al. introduce a model for the Epps effect which is based on the phenomenon of lagged correlations [19].

However, certainly miscellaneous mechanisms are contributing to the Epps effect. Thus our approach is different. First, we will introduce a simple model which offers an explanation for the statistical part of the Epps effect, based on a central assumption of an underlying time series. Secondly, based on that model, we will present an estimator, with which these effects can be compensated. Finally, we will quantify the impact of this phenomenon on the Epps effect in recent empirical data and show that it can be a major cause for the Epps effect, especially when looking at less frequently traded securities.

This paper is organized as followed: In section 2, we develop the model for correlations in asynchronous time series. Within the model, we observe a decay of correlations towards smaller return intervals, similar to the Epps effect. We then derive the method to compensate this phenomenon. In section 3, this method is applied to recent empirical data to estimate the impact of the observed effect on the Epps effect. We discuss the results in section 4.

*Corresponding author. Tel.: +49 203 379 4727; Fax: +49 203 379 4732.

Email address: michael@mueninx.com (Michael C. Münnix)

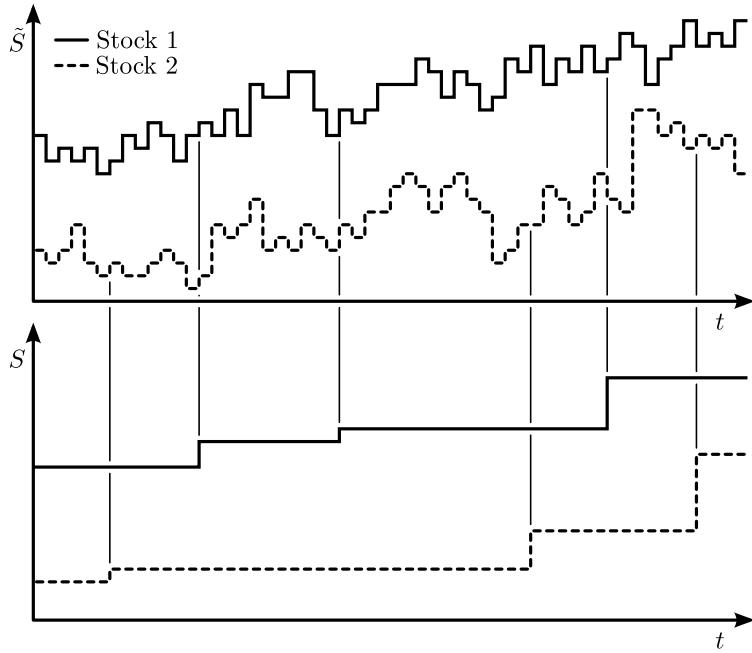


Figure 1: Illustration of the model for asynchronous trading times of two stocks. Shown above are the prices \tilde{S} on the underlying timescale. The “sampling” of these prices \tilde{S} to prices S on simulated trading times are shown below.

2. Statistical effects in asynchronous time series

In section 2.1, we set up our model and develop a compensation formalism for asynchrony effects in section 2.2.

2.1. Model

The central assumption of our model is the existence of an underlying non-lagged time series of prices. The assumption of a finer [19] or even continuous [11, 20, 18] underlying timescale is a common approach in the estimation of correlations. This approach is also intuitive, as most stocks are traded at several stock exchanges simultaneously.

To simulate asynchrony effects we generate an underlying correlated time series using the *Capital Asset Pricing Model* (CAPM) [21], which is also known as Noh’s model [22] in physics,

$$\tilde{r}^{(i)}(t) = \sqrt{c} \eta(t) + \sqrt{1-c} \varepsilon^{(i)}(t), \quad (1)$$

where $\tilde{r}^{(i)}$ stands for the relative price change, the so-called *return* of the i -th stock and c is the correlation coefficient. The random variables η and $\varepsilon^{(i)}$ are taken from a compound distribution as observed on market data by Gopikrishnan et. al. with power-law tails and a central Levy distribution (for details see Ref. [23]). We have chosen this approach to keep our model initially as simple as possible. We note, however, that return time series can also be autocorrelated. While first order autocorrelations are in this context insignificantly small [24], second order autocorrelations or “volatility clustering” represent a strong characteristic of return time series and led to the development of autoregressive models, such as GARCH [25, 26]. For this reason we also test our compensation in a more realistic setup against a GARCH(1,1) generated time series of underlying returns, given by

$$\tilde{r}^{(i)}(t) = \sigma^{(i)}(t) \left(\sqrt{c} \eta(t) + \sqrt{1-c} \varepsilon^{(i)}(t) \right) \quad (2)$$

with

$$(\sigma^{(i)}(t))^2 = \alpha_0 + \alpha_1 (\tilde{r}^{(i)}(t-1))^2 + \beta_1 (\sigma^{(i)}(t-1))^2. \quad (3)$$

The initial parameters of the GARCH process have been chosen as $\alpha_0 = 2.4 \cdot 10^{-4}$, $\alpha_1 = 0.15$ and $\beta_1 = 0.84$.

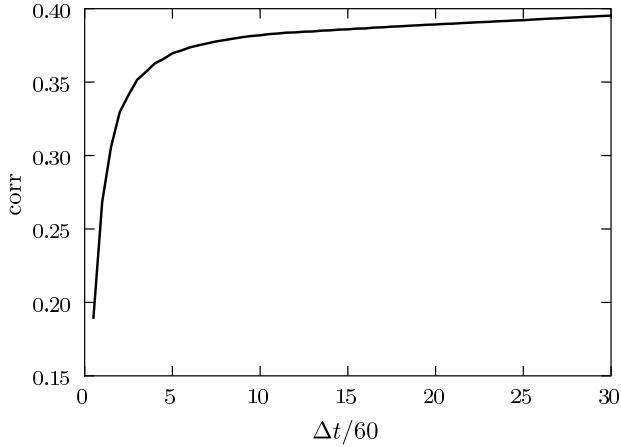


Figure 2: Scaling behavior of the correlation coefficient on a simulated asynchronous time series. The length of the underlying time series was set to $7.2 \cdot 10^6$ ($c = 0.4$, $\mu^{(1)} = 15$ and $\mu^{(2)} = 25$). The correlation coefficient was calculated on return-intervals from 60 data points (corresponding to 1 minute) to 1800 data points (corresponding to 30 minutes).

Two return time series $\tilde{r}^{(1)}$ and $\tilde{r}^{(2)}$ are generated representing two correlated stocks. The lengths of these underlying time series are chosen as $7.2 \cdot 10^6$, $1.44 \cdot 10^6$ and $7.2 \cdot 10^5$ corresponding to a return interval $\Delta\tilde{t}$ on the underlying timescale of 1, 5 and 10 seconds during 1 trading year.

Using these returns and an arbitrary starting price, the underlying price series $\tilde{S}^{(1)}$ and $\tilde{S}^{(2)}$ are calculated implying a geometric Brownian motion with zero drift and a standard deviation of 10^{-3} per time step. To model the asynchronous trade processes, these prices are sampled independently using exponentially distributed waiting times with average values typical for the stock market (see Fig. 1). In the following example, we choose the average waiting times as $\mu^{(1)} = 15$ and $\mu^{(2)} = 25$ (equivalent to seconds in this example), while the underlying time series were correlated with $c = 0.4$. On the resulting ‘macroscopic’ time series, the return between two points in time (of the i -th stock) can be calculated as

$$r_{\Delta t}^{(i)}(t) = \frac{S^{(i)}(t + \Delta t) - S^{(i)}(t)}{S^{(i)}(t)}, \quad (4)$$

where $S^{(i)}(t)$ denotes the price at time t and Δt is the return interval. Between these return time series, we now calculate the correlation coefficient,

$$\text{corr}(r_{\Delta t}^{(i)}, r_{\Delta t}^{(j)}) = \frac{\langle r_{\Delta t}^{(i)} r_{\Delta t}^{(j)} \rangle - \langle r_{\Delta t}^{(i)} \rangle \langle r_{\Delta t}^{(j)} \rangle}{\sigma_{\Delta t}^{(i)} \sigma_{\Delta t}^{(j)}}, \quad (5)$$

where $\langle \dots \rangle$ denotes the mean value of a time series with length T and where σ refers to the standard deviation of the same time series. We note that we refer to the whole time series of returns $r_{\Delta t}^{(i)}$ when the argument (t) is omitted.

When calculating the correlation of returns of the sampled time series \tilde{S} using different return intervals Δt , the correlation coefficient scales down as shown in Fig 2. This behavior is very similar to the Epps effect in empirical data. It occurs only because of the asynchrony of the trading times. As this behavior is already observed in this simple setting, we are able to derive a method to compensate it, as the following demonstrates.

2.2. Compensation

The basic idea of this approach is the following: Due to the asynchrony, each term of the correlation coefficient can be divided into a part which contributes to the correlation and a part which is uncorrelated and therefore lowers the correlation coefficient.

According to the model assumption, the price change during Δt is based on price changes on an underlying ‘microscopic’ timescale. Thus, the return can also be expressed as a sum of the underlying returns,

$$r^{(i)}(t) = \sum_{j=0}^{N_{\Delta t}^{(i)}(t)} \tilde{r}^{(i)}(\gamma^{(i)}(t) + j\Delta\tilde{t}) . \quad (6)$$

Here $\tilde{r}^{(i)}(t_i)$ is the return related to $S(t)$ on the underlying time scale of non-overlapping intervals $\Delta\tilde{t}$ (e.g. 1 second) given by

$$\tilde{r}^{(i)}(t + j\Delta\tilde{t}) = \frac{\tilde{S}(t + (j+1)\Delta\tilde{t}) - \tilde{S}(t + j\Delta\tilde{t})}{S(t)} . \quad (7)$$

The quantity $\gamma^{(i)}(t)$ in equation (6) represents the time of the last trade of the i -th stock at time t ,

$$\gamma^{(i)}(t) = \max(t_{\text{trade}}^{(i)}) \Big|_{t_{\text{trade}}^{(i)} \leq t} . \quad (8)$$

When calculating the return for the interval $[t, t + \Delta t]$ of two stocks, the actual price at t and $t + \Delta t$ is generally in the past, more precisely at $\gamma^{(1)}(t)$, $\gamma^{(2)}(t)$ and $\gamma^{(1)}(t + \Delta t)$, $\gamma^{(2)}(t + \Delta t)$. These trading times are distinct for each stock, therefore only a fraction of the underlying prices processed by the return is correlated. The number of terms $N_{\Delta t}^{(i)}$ of the sum in equation (6) is given by

$$N_{\Delta t}^{(i)}(t) = \frac{(\gamma^{(i)}(t + \Delta t) - \gamma^{(i)}(t))}{\Delta\tilde{t}} . \quad (9)$$

We normalize the returns to zero mean and unit variance and indicate them as g and \tilde{g} :

$$g_{\Delta t}^{(i)}(t) = \frac{r_{\Delta t}^{(i)}(t) - \langle r_{\Delta t}^{(i)} \rangle}{\sigma_{\Delta t}^{(i)}}, \quad \tilde{g}^{(i)}(t) = \frac{\tilde{r}^{(i)}(t) - \langle \tilde{r}^{(i)} \rangle}{\tilde{\sigma}^{(i)}} . \quad (10)$$

In this context, the relation of the returns on both time scales in equation (6) changes to

$$g_{\Delta t}^{(i)}(t) = \sqrt{\frac{\Delta\tilde{t}}{\Delta t}} \sum_{j=0}^{N_{\Delta t}^{(i)}(t)} \tilde{g}^{(i)}(\gamma^{(i)}(t) + j\Delta\tilde{t}) - \frac{\langle \tilde{r}^{(i)} \rangle \left(\frac{\Delta t}{\Delta\tilde{t}} - N_{\Delta t}^{(i)}(t) \right)}{\sigma_{\Delta t}^{(i)}} , \quad (11)$$

as worked out in appendix Appendix A. When using normalized returns, the correlation coefficient of two return time series $r_{\Delta t}^{(1)}$ and $r_{\Delta t}^{(2)}$ (see equation (5)) simplifies to

$$\text{corr}(r_{\Delta t}^{(1)}, r_{\Delta t}^{(2)}) = \text{corr}(g_{\Delta t}^{(1)}, g_{\Delta t}^{(2)}) = \frac{1}{T} \sum_{j=0}^T g_{\Delta t}^{(1)}(t_j) g_{\Delta t}^{(2)}(t_j) . \quad (12)$$

As the mean value over T of the second term from equation (11) is equal to zero, we obtain in terms of the underlying time series

$$\text{corr}(r_{\Delta t}^{(1)}, r_{\Delta t}^{(2)}) = \frac{1}{T} \sum_{j=0}^T \left(\sum_{k=0}^{N_{\Delta t}^{(1)}(t_j)} \tilde{g}^{(1)}(\gamma^{(1)}(t_j) + k\Delta\tilde{t}) \sum_{l=0}^{N_{\Delta t}^{(2)}(t_j)} \tilde{g}^{(2)}(\gamma^{(2)}(t_j) + l\Delta\tilde{t}) \right) \frac{\Delta\tilde{t}}{\Delta t} . \quad (13)$$

As illustrated in Fig. 3, only a subset of the underlying prices \tilde{S} of two prices S share an overlapping time-interval. Because of this “overlap” only a certain amount $\tilde{N}_{\Delta t}(t)$ of the underlying returns is correlated, namely

$$\tilde{N}_{\Delta t}(t) = \frac{\Delta t_o(t)}{\Delta\tilde{t}} \quad (14)$$

with $\Delta t_o(t)$ being the time interval of the actual overlap,

$$\Delta t_o(t) = \min(\gamma^{(1)}(t + \Delta t), \gamma^{(2)}(t + \Delta t)) - \max(\gamma^{(1)}(t), \gamma^{(2)}(t)) . \quad (15)$$

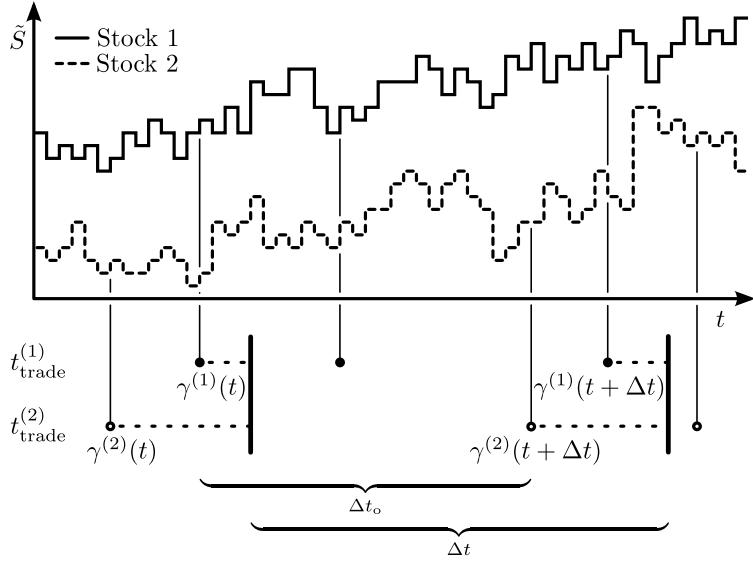


Figure 3: Illustration of the overlap Δt_o .

Each sum can be split up into $N_{\Delta t}^{(i)} - \bar{N}$ terms that are uncorrelated and \bar{N} that are correlated. Thus, equation (13) can be written as:

$$\text{corr}(r_{\Delta t}^{(1)}, r_{\Delta t}^{(2)}) = \frac{1}{T} \sum_{j=0}^T \left(\left(\underbrace{\sum_{k=\bar{N}_{\Delta t}(t_j)+1}^{N_{\Delta t}^{(1)}(t_j)-\bar{N}_{\Delta t}(t_j)} \tilde{g}^{(1)}(t_k)}_{\text{async.}} + \underbrace{\sum_{\bar{k}=0}^{\bar{N}_{\Delta t}(t_j)} \tilde{g}^{(1)}(t_{\bar{k}})}_{\text{sync.}} \right) \times \left(\underbrace{\sum_{l=\bar{N}_{\Delta t}(t_j)+1}^{N_{\Delta t}^{(2)}(t_j)-\bar{N}_{\Delta t}(t_j)} \tilde{g}^{(2)}(t_l)}_{\text{async.}} + \underbrace{\sum_{\bar{l}=0}^{\bar{N}_{\Delta t}(t_j)} \tilde{g}^{(2)}(t_{\bar{l}})}_{\text{sync.}} \right) \frac{\Delta \tilde{t}}{\Delta t} \right), \quad (16)$$

where only the sums of synchronous returns are correlated among each other. In this notation, the underlying time series is indexed as $[\tilde{r}^{(i)}(t_0), \tilde{r}^{(i)}(t_1), \dots, \tilde{r}^{(i)}(t_{N_{\Delta t}^{(i)}})]$, where the returns from t_0 to $t_{\bar{N}_{\Delta t}}$ are corresponding to the overlap. When expanding the product, the non-correlated returns converge to zero due to the outer average

$$\begin{aligned} \text{corr}(r_{\Delta t}^{(1)}, r_{\Delta t}^{(2)}) &= \frac{1}{T} \sum_{j=0}^T \left(\left(\underbrace{\sum_{k=0}^{\bar{N}_{\Delta t}(t_j)} \tilde{g}^{(1)}(t_k) \tilde{g}^{(2)}(t_k)}_{\bar{N}_{\Delta t}(t) \text{corr}_{t_j}(\tilde{r}_1, \tilde{r}_2)} + \underbrace{\dots}_{0} \right) \frac{\Delta \tilde{t}}{\Delta t} \right) \\ &= \frac{1}{T} \sum_{j=0}^T \text{corr}_{t_j}(\tilde{g}^{(1)}, \tilde{g}^{(2)}) \frac{\bar{N}_{\Delta t}(t_j) \Delta \tilde{t}}{\Delta t} \\ &= \frac{1}{T} \sum_{j=0}^T \text{corr}_{t_j}(\tilde{g}^{(1)}, \tilde{g}^{(2)}) \frac{\Delta t_o(t_j)}{\Delta t}, \end{aligned} \quad (17)$$

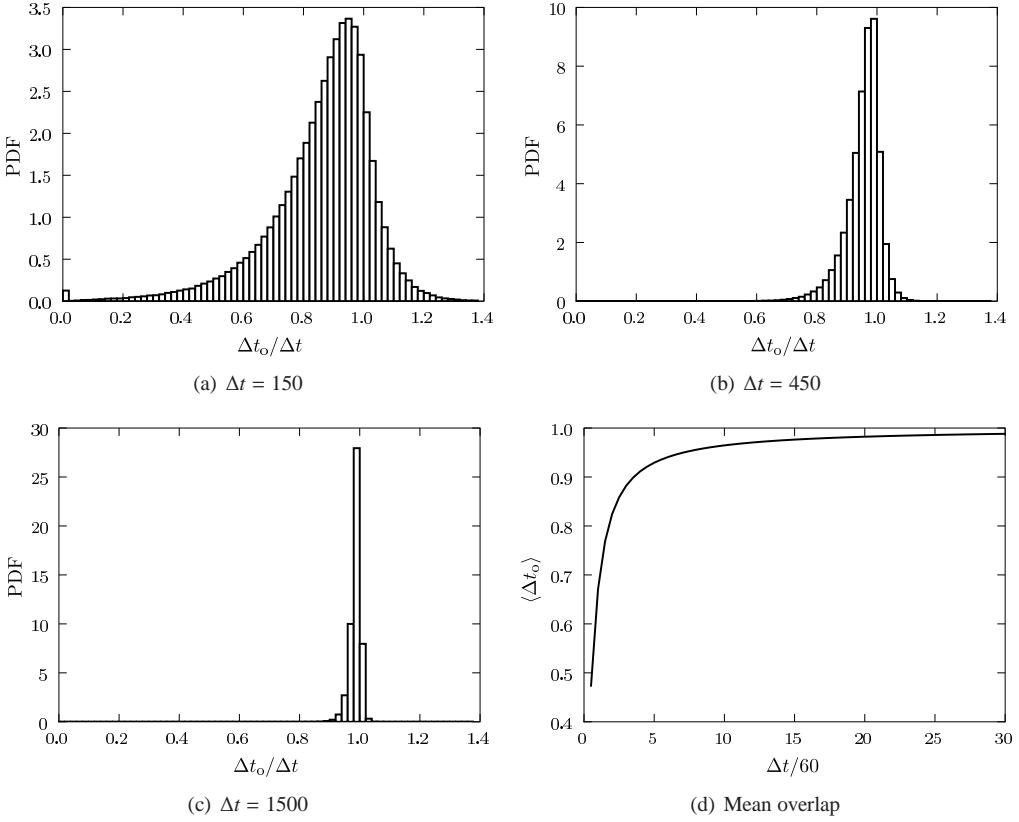


Figure 4: Distribution of the overlaps for simulated return intervals $\Delta t = 150$ (a), 450 (b) and 1500 (c) data points (corresponding to 2.5 , 7.5 and 25 minutes). Towards larger return-intervals, the distribution sharpens, as well as its mean converges to 1 (d).

where corr_t represents the correlation of the underlying returns corresponding to the interval $[t, t + \Delta t]$.

$\Delta t_o(t)/\Delta t$ is the fractional overlap of the corresponding return interval. The fractional overlap does not depend on the actual timescale of the underlying time series. As equation (17) clearly shows, the correlation coefficient of the synchronous part of the return time series is multiplied by the fractional overlap. Hence, this effect can be compensated by

$$\text{corr}_{\text{corrected}}(r_{\Delta t}^{(1)}, r_{\Delta t}^{(2)}) = \frac{1}{T} \sum_{j=0}^T g_{\Delta t}^{(1)}(t_j) g_{\Delta t}^{(2)}(t_j) \frac{\Delta t}{\Delta t_o(t_j)}. \quad (18)$$

The dashed line in Fig. 5 represents the asynchrony-compensated correlation within our simulation. It turns out that there is a remaining effect that still causes a downscaling of the correlation coefficient for very small return intervals. This behavior occurs when the price of either of the stocks did not change during the return-interval and therefore the corresponding return equals zero. Of course, this event becomes more probable on smaller return intervals Δt . It corresponds to the small peak at $\Delta t_o = 0$ in Fig. 4(a). This remaining downscaling coincides with the cumulative estimator described by Hayashi and Yoshida [11]. It can also be expressed in the formalism used here. It reads

$$\text{corr}(r_{\Delta t}^{(1)}, r_{\Delta t}^{(2)}) \Big|_{(\gamma_1^{(1)}(t) \neq \gamma^{(1)}(t + \Delta t)) \wedge (\gamma^{(2)}(t) \neq \gamma^{(2)}(t + \Delta t))}. \quad (19)$$

Therefore, when combining both estimators, and thus only regarding returns with overlapping time intervals, the remaining scaling behavior for very small returns can be compensated as well.

As displayed in Fig. 4, the overlap can also be larger than the actual return interval, implying that terms with such overlaps are corrected downwards. Therefore the compensation can amplify a specific term of the correlation coefficient as well as it can attenuate it.

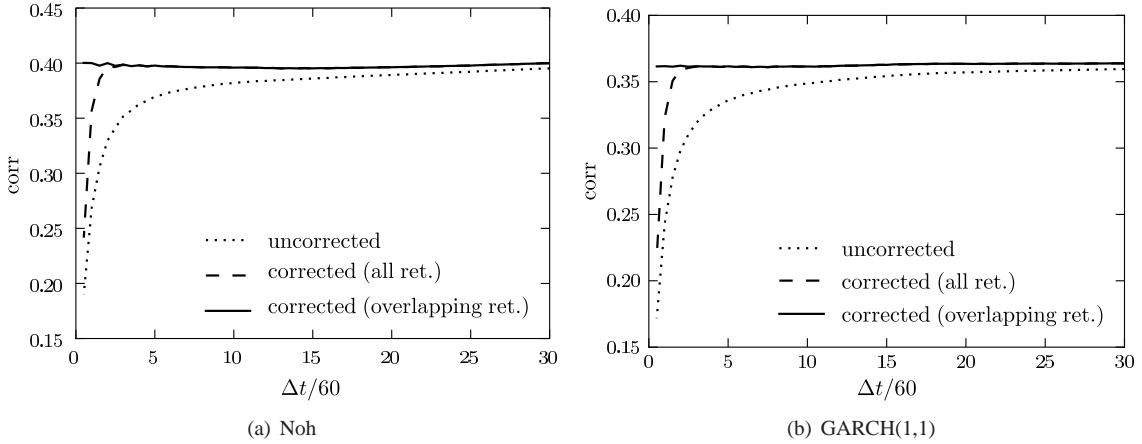


Figure 5: Compensation of asynchrony effects within the model with different approaches in the generation of the underlying time series. The dashed line represents the correlation coefficient, which is corrected by the overlap. The solid line regards in addition only returns, in which time intervals trades occurred.

3. Application to market data

Certainly, many aspects contribute to the Epps effect. Our present aim is to quantify the part, which is caused by the asynchrony of the time series.

It is difficult to isolate the Epps Effect on single stock pairs, as it can superimpose with other effects leading to other characteristics of the correlation coefficient than expected according to the Epps effect. A common approach on this topic is to pick the pairs of stocks, which show a distinctive Epps effect and focus the analysis to these pairs [3, 19, 12]. In the following, we would like to take a different approach:

We classify two ensembles of stock pairs. After compensating the asynchrony effect for each pair, we build the average for the ensemble. We also plot the error bars representing the double standard deviation 2σ . By this method, we can show the scope of the asynchrony model and identify regions, in which other effects dominate. All data was extracted from the NYSE's TAQ database for the year 2007 [27].

The first ensemble consists of stock pairs which provide the most stable correlation. Thereby we want to suppress those effects which are caused by a change in the correlation during the period in which the correlation coefficient is calculated. This ensemble represents ideal test conditions for the asynchrony compensation. To identify those stock pairs with a stable correlation, we calculate the correlation coefficient of 30 daily returns. After shifting this window in 1-day intervals through the year, we calculate the variance of the obtained correlation coefficients (var_{corr}). Then we identify the five stocks providing the smallest variance for each *Global Industry Classification System* (GICS) branch of the *Standard & Poor's* (S&P) 500 index. This results in an ensemble of 50 stocks as shown in table B.1, appendix Appendix B.

As the correlation structure of stocks can be non-stationary, we also evaluate the asynchrony compensation without the restriction to stable correlations. For this purpose, we select a second ensemble consisting of 5 stock pairs of each GICS branch of the S&P 500 index, whose daily returns are providing the strongest correlation during the year 2007. These stocks include highly non-stationary correlations as indicated in table B.2, appendix Appendix B (row “ var_{corr} ”).

Fig. 6 shows the ensemble average of the correlation coefficient and the asynchrony-compensated correlation coefficient for both ensembles in 2007 (250 trading days). Before averaging, the correlation coefficients for each stock have been normalized to the value at a return interval $\Delta t = 40$ minutes.

When looking at the whole ensemble we discover that the asynchrony has a pronounced impact on the Epps effect. The asynchrony effect seems to be the dominating cause for the Epps effect on return intervals down to approximately 10 minutes, where the remaining Epps effect is on average less than 3% of the correlation coefficient’s saturation value at large return intervals. For smaller return intervals, other effects dominate, e.g. a lag between the time series of two

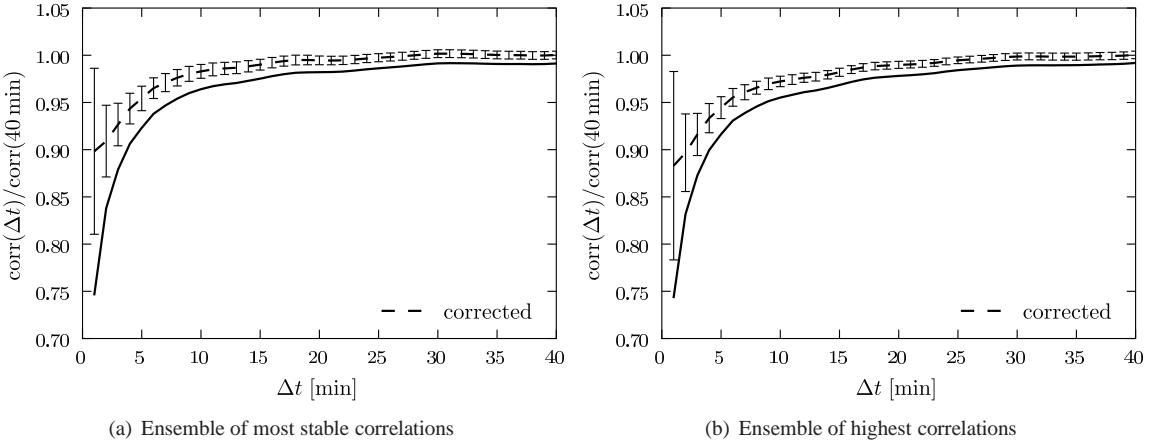


Figure 6: Asynchrony-compensated correlations of two ensembles. The data has been normalized to its value at 40 minutes. The error bars represent the double standard deviation.

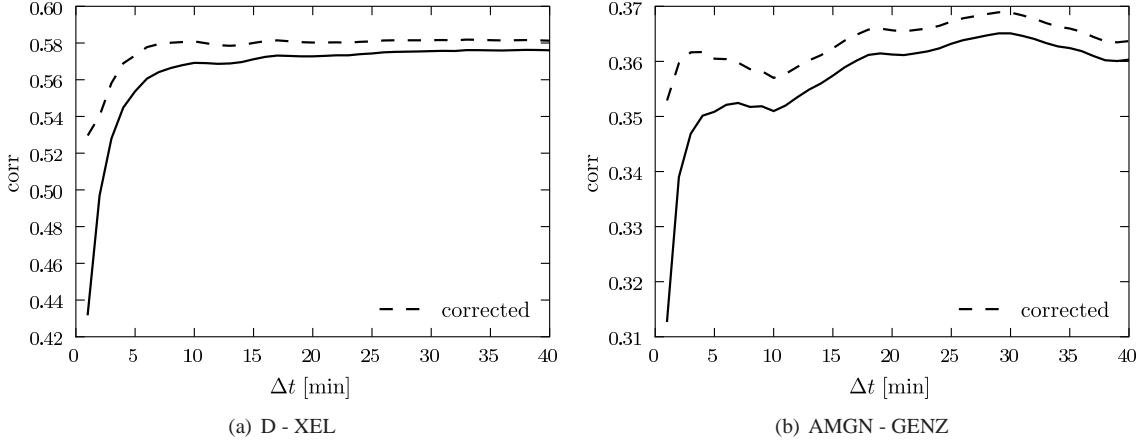


Figure 7: Asynchrony-compensated correlations between Dominion Resources, Inc. (D) - Xcel Energy Inc. (XEL) and Amgen Inc. (AMGN) - Genzyme Corp. (GENZ)

stocks, as recent study indicates [19].

However, the ensemble consists also of stocks which are very frequently traded, providing a very short average waiting time which results in a fractional overlap $\Delta_{t_0}(t)/\Delta t$ close to unity. Evidently the presented compensation only has a small impact on the correlation estimation of these stocks, as they are so frequently traded that their time series can almost be described as continuous. Naturally the presented compensation works best for less frequently traded stocks, as they actually show an asynchronous behavior. Fig. 7 provides two examples of stock pairs for which the asynchrony of time series is a major effect. While Fig. 7(a) shows a “clean” Epps effect, Fig. 7(b) shows an Epps effect which is superimposed with other phenomena.

Of course, within the statistical ensemble stock pairs can be found that either do not show an Epps effect or that are so infrequently traded that the assumption of an underlying timeline seems to be unreasonable. Even though the assumption of an underlying time series is a common and intuitive approach on this topic, it may not be valid for very infrequently traded stocks.

When looking at single stock pairs, it turns out that the asynchrony-compensation works well, if a distinguished Epps effect is found. In case of adopting the presented method as a black box model without looking at the scaling

behavior of the correlation coefficient, we believe that a return interval of 5 minutes represents a good lower bound for the scope of this method.

4. Conclusion

We presented a model for the scaling behavior of financial correlations due to the asynchrony of the time series. This purely statistical effect can be compensated. Furthermore, we applied this compensation to market data under the assumption of an underlying time series with non-lagged correlations. We quantified the influence of the asynchrony on the overall decay of the correlation coefficient towards small return intervals, which is known as the Epps effect. The results clearly demonstrate that the asynchrony can have a huge impact on the Epps effect. It rather can be the dominating cause for less frequently traded stocks. The main advantage of our method is that no parameters or adjustments are necessary, since it is based on the trading times only.

In our empirical study, the asynchrony-compensation allowed us to recover the correlation coefficient for return intervals down to 10 minutes. At this return interval, the remaining Epps effect is on average less than 3% of the correlation coefficient's saturation value at large return intervals. We also demonstrated that the presented method holds for non-stationary correlated time series. The accurate calculation of correlations is of major importance for risk management. To keep the estimation error small, a long time series of returns is required. Yet at the same time, the time series should not reach too far into the past. The latter is important because the correlation structure can be highly dynamic, as the dramatic events of autumn 2008 prove. Applying our method to intraday data allows to choose smaller return intervals and hence provides improved statistical significance of the correlations for the same time horizon.

Acknowledgements

The authors thank J. Hämerling and V. Osipov for fruitful discussion. M.C.M. acknowledges financial support from Studienstiftung des deutschen Volkes.

Appendix A. Relation between $g_{\Delta t}^{(i)}$ and $\tilde{g}^{(i)}$

We defined the normalized returns as

$$g_{\Delta t}^{(i)}(t) = \frac{r_{\Delta t}^{(i)}(t) - \langle r_{\Delta t}^{(i)} \rangle}{\sqrt{\text{Var}(r_{\Delta t}^{(i)})}} \quad (\text{A.1})$$

$$\tilde{g}^{(i)}(t) = \frac{\tilde{r}^{(i)}(t) - \langle \tilde{r}^{(i)} \rangle}{\sqrt{\text{Var}(\tilde{r}^{(i)})}}, \quad (\text{A.2})$$

where $\text{Var}(\dots)$ refers to the variance of a time series. Inserting the return, expressed through the underlying time series,

$$r^{(i)}(t) = \sum_{j=0}^{N_{\Delta t}^{(i)}(t)} \tilde{r}^{(i)}(\gamma^{(i)}(t) + j\Delta \tilde{t}), \quad (\text{A.3})$$

in equation (A.1), results in

$$g_{\Delta t}^{(i)}(t) = \frac{\sum_{j=0}^{N_{\Delta t}^{(i)}(t)} (\tilde{r}^{(i)}(\gamma^{(i)}(t) + j\Delta \tilde{t}) - \langle r_{\Delta t}^{(i)} \rangle)}{\sqrt{\text{Var}(r_{\Delta t}^{(i)})}} \quad (\text{A.4})$$

$$= \frac{\sqrt{\text{Var}(\tilde{r}^{(i)})} \sum_{j=0}^{N_{\Delta t}^{(i)}(t)} (\tilde{g}^{(i)}(\gamma^{(i)}(t) + j\Delta \tilde{t}) - \langle r_{\Delta t}^{(i)} \rangle + N_{\Delta t}^{(i)}(t) \langle \tilde{r}^{(i)} \rangle)}{\sqrt{\text{Var}(r_{\Delta t}^{(i)})}}. \quad (\text{A.5})$$

In equation (A.5), (A.2) was used to express the underlying returns $\tilde{r}^{(i)}$.

The mean values and variance are additive, which leads to

$$\langle r_{\Delta t}^{(i)} \rangle = \langle N_{\Delta t}^{(i)}(t) \rangle \langle \tilde{r}^{(i)} \rangle, \quad \text{Var}(r_{\Delta t}^{(i)}) = \langle N_{\Delta t}^{(i)}(t) \rangle \text{Var}(\tilde{r}^{(i)}). \quad (\text{A.6})$$

Therefore, we obtain

$$g_{\Delta t}^{(i)}(t) = \frac{1}{\sqrt{\langle N_{\Delta t}^{(i)}(t) \rangle}} \sum_{j=0}^{N_{\Delta t}^{(i)}(t)} \tilde{g}^{(i)}(\gamma^{(i)}(t) + j\Delta\tilde{t}) - \frac{\langle \tilde{r}^{(i)} \rangle (\langle N_{\Delta t}^{(i)}(t) \rangle - N_{\Delta t}^{(i)}(t))}{\sqrt{\text{Var}(r_{\Delta t}^{(i)})}}. \quad (\text{A.7})$$

As the average time interval per return converges to Δt , the mean number of underlying price changes $\langle N_{\Delta t}^{(i)}(t) \rangle$ is given by $\Delta t/\Delta\tilde{t}$. Thus, we arrive at

$$g_{\Delta t}^{(i)}(t) = \sqrt{\frac{\Delta\tilde{t}}{\Delta t}} \sum_{j=0}^{N_{\Delta t}^{(i)}(t)} \tilde{g}^{(i)}(\gamma^{(i)}(t) + j\Delta\tilde{t}) - \frac{\langle \tilde{r}^{(i)} \rangle \left(\frac{\Delta t}{\Delta\tilde{t}} - N_{\Delta t}^{(i)}(t)\right)}{\sqrt{\text{Var}(r_{\Delta t}^{(i)})}}, \quad (\text{A.8})$$

which is equation (11).

Appendix B. Stock ensembles

Table B.1: Top 5 five stock pairs with the most stable correlation from each GICS branch of the S&P 500 index.

| GICS Branch | Symbol | Name | Stock 1 | | | Stock 2 | | | corr | varcorr |
|----------------------------|--------|---------------------------------|----------------|------------|--------|----------------------------|----------------|------------|-------|---------|
| | | | Stock Exchange | Volume | Symbol | Name | Stock Exchange | Volume | | |
| Consumer Discretionary | AMZN | Amazon Corp. | NASDAQ | 1215700 | SBUX | Starbucks Corp. | NASDAQ | 1287200 | 0.80 | 0.45 |
| | APOL | Apollo Group | NASDAQ | 332200 | SHLD | Sears Holdings Corporation | NASDAQ | 317900 | 0.28 | 0.57 |
| | AMZN | Amazon Corp. | NASDAQ | 1215700 | SPLS | Staples Inc. | NASDAQ | 674500 | 0.80 | 0.62 |
| | AMZN | Amazon Corp. | NASDAQ | 1215700 | CMCSA | Comcast Corp. | NASDAQ | 910000 | 0.69 | 0.64 |
| | EXPE | Expedia Inc. | NASDAQ | 883300 | SHLD | Sears Holdings Corporation | NASDAQ | 317900 | 0.43 | 0.66 |
| Consumer Staples | COST | Costco Co. | NASDAQ | 237500 | PG | Procter & Gamble | NYSE | 1396077700 | 0.05 | 0.71 |
| | COST | Costco Co. | NASDAQ | 237500 | CVS | CVS Caremark Corp. | NYSE | 1454768200 | 0.09 | 0.77 |
| | KO | Coca Cola Co. | NYSE | 1152559200 | COST | Costco Co. | NASDAQ | 237500 | 0.04 | 0.79 |
| | MO | Altria Group Inc. | NYSE | 1312801100 | CCE | Coca-Cola Enterprises | NYSE | 335449400 | 0.29 | 0.79 |
| | CCE | Coca-Cola Enterprises | NYSE | 335449400 | KFT | Kraft Foods Inc-A | NYSE | 1438368700 | 0.32 | 0.80 |
| Energy | EP | El Paso Corp. | NYSE | 697103700 | SE | Spectra Energy Corp. | NYSE | 331523100 | 0.34 | 0.84 |
| | CVX | Chevron Corp. | NYSE | 1271849400 | SE | Spectra Energy Corp. | NYSE | 331523100 | 0.26 | 0.85 |
| | HES | Hess Corporation | NYSE | 427802700 | SE | Spectra Energy Corp. | NYSE | 331523100 | 0.23 | 0.85 |
| | MUR | Murphy Oil | NYSE | 223916300 | SE | Spectra Energy Corp. | NYSE | 331523100 | 0.25 | 0.86 |
| | SII | Smith International | NYSE | 370893400 | SE | Spectra Energy Corp. | NYSE | 331523100 | 0.32 | 0.86 |
| Financials | SCHW | Charles Schwab | NASDAQ | 1445500 | ETFC | E*Trade Financial Corp. | NASDAQ | 1391800 | 0.67 | 0.44 |
| | ETFC | E*Trade Financial Corp. | NASDAQ | 1391800 | FITB | Fifth Third Bancorp | NASDAQ | 218000 | 0.29 | 0.58 |
| | SCHW | Charles Schwab | NASDAQ | 1445500 | FITB | Fifth Third Bancorp | NASDAQ | 218000 | 0.36 | 0.58 |
| | SCHW | Charles Schwab | NASDAQ | 1445500 | HCBK | Hudson City Bancorp | NASDAQ | 686900 | 0.54 | 0.60 |
| | ACAS | American Capital Strategies Ltd | NASDAQ | 207200 | SCHW | Charles Schwab | NASDAQ | 1445500 | 0.50 | 0.60 |
| Health Care | CELG | Celgene Corp. | NASDAQ | 619200 | ESRX | Express Scripts | NASDAQ | 998400 | 0.65 | 0.47 |
| | AMGN | Amgen | NASDAQ | 813900 | CELG | Celgene Corp. | NASDAQ | 619200 | 0.48 | 0.50 |
| | AMGN | Amgen | NASDAQ | 813900 | BIIB | BIOGEN IDEC Inc. | NASDAQ | 381800 | 0.48 | 0.52 |
| | CELG | Celgene Corp. | NASDAQ | 619200 | THC | Tenet Healthcare Corp. | NYSE | 805228900 | -0.23 | 0.53 |
| | AMGN | Amgen | NASDAQ | 813900 | GENZ | Genzyme Corp. | NASDAQ | 242900 | 0.57 | 0.53 |
| Industrials | GE | General Electric | NYSE | 4303823300 | LUV | Southwest Airlines | NYSE | 862775700 | 0.53 | 0.73 |
| | MMM | 3M Company | NYSE | 549124400 | CBE | Cooper Industries Ltd. | NYSE | 175911300 | 0.24 | 0.73 |
| | CBE | Cooper Industries Ltd. | NYSE | 175911300 | GWV | Grainger (W.W.) Inc. | NYSE | 95324400 | 0.15 | 0.73 |
| | CBE | Cooper Industries Ltd. | NYSE | 175911300 | GR | Goodrich Corporation | NYSE | 144177800 | 0.15 | 0.75 |
| | CBE | Cooper Industries Ltd. | NYSE | 175911300 | FLR | Fluor Corp. (New) | NYSE | 171713100 | 0.10 | 0.76 |
| Information Technology | AAPL | Apple Inc. | NASDAQ | 4627500 | INTC | Intel Corp. | NASDAQ | 5529100 | 0.84 | 0.19 |
| | AAPL | Apple Inc. | NASDAQ | 4627500 | CSCO | Cisco Systems | NASDAQ | 4886800 | 0.76 | 0.24 |
| | AAPL | Apple Inc. | NASDAQ | 4627500 | YHOO | Yahoo Inc. | NASDAQ | 2609700 | 0.71 | 0.25 |
| | AAPL | Apple Inc. | NASDAQ | 4627500 | ORCL | Oracle Corp. | NASDAQ | 1731900 | 0.78 | 0.25 |
| | AAPL | Apple Inc. | NASDAQ | 4627500 | EBAV | eBay Inc. | NASDAQ | 1056100 | 0.73 | 0.26 |
| Materials | MON | Monsanto Co. | NYSE | 479605700 | SEE | Sealed Air Corp.(New) | NYSE | 126716200 | 0.06 | 0.52 |
| | FCX | Freeport-McMoran Cp & Gld | NYSE | 1058215000 | SIAL | Sigma-Aldrich | NASDAQ | 133800 | 0.05 | 0.54 |
| | ECL | Ecolab Inc. | NYSE | 163404500 | SEE | Sealed Air Corp.(New) | NYSE | 126716200 | 0.26 | 0.60 |
| | ATI | Allegheny Technologies Inc | NYSE | 269746100 | SEE | Sealed Air Corp.(New) | NYSE | 126716200 | 0.10 | 0.63 |
| | PX | Praxair Inc. | NYSE | 245761900 | SEE | Sealed Air Corp.(New) | NYSE | 126716200 | 0.12 | 0.65 |
| Telecommunication Services | Q | Qwest Communications Int | NYSE | 1623807700 | S | Sprint Nextel Corp. | NYSE | 2044634000 | 0.52 | 0.84 |
| | Q | Qwest Communications Int | NYSE | 1623807700 | VZ | Verizon Communications | NYSE | 1472335800 | 0.49 | 0.86 |
| | S | Sprint Nextel Corp. | NYSE | 2044634000 | VZ | Verizon Communications | NYSE | 1472335800 | 0.49 | 0.87 |
| | AMT | American Tower Corp. | NYSE | 387199400 | Q | Qwest Communications Int | NYSE | 1623807700 | 0.26 | 0.97 |
| | AMT | American Tower Corp. | NYSE | 387199400 | WIN | Windstream Corporation | NYSE | 400634200 | 0.10 | 0.99 |
| Utilities | DUK | Duke Energy | NYSE | 902519200 | DYN | Dynegy Inc. | NYSE | 702035600 | 0.54 | 0.74 |
| | CMS | CMS Energy | NYSE | 264225200 | DYN | Dynegy Inc. | NYSE | 702035600 | 0.48 | 0.79 |
| | CMS | CMS Energy | NYSE | 264225200 | DUK | Duke Energy | NYSE | 902519200 | 0.39 | 0.81 |
| | AES | AES Corp. | NYSE | 556049300 | CMS | CMS Energy | NYSE | 264225200 | 0.31 | 0.84 |
| | CNP | CenterPoint Energy | NYSE | 359757800 | DUK | Duke Energy | NYSE | 902519200 | 0.33 | 0.84 |

Table B.2: Top 5 five stock pairs with the highest correlation from each GICS branch of the S&P 500 index.

| GICS Branch | Symbol | Name | Stock 1 | | | Stock 2 | | | corr | varcorr |
|----------------------------|--------|---------------------------|----------------|------------|--------|----------------------------------|----------------|------------|------|---------|
| | | | Stock Exchange | Volume | Symbol | Name | Stock Exchange | Volume | | |
| Consumer Discretionary | APOL | Apollo Group | NASDAQ | 332200 | SPLS | Staples Inc. | NASDAQ | 674500 | 0.77 | 1.08 |
| | BBBY | Bed Bath & Beyond | NASDAQ | 233000 | SPLS | Staples Inc. | NASDAQ | 674500 | 0.79 | 1.08 |
| | AMZN | Amazon Corp. | NASDAQ | 1215700 | SPLS | Staples Inc. | NASDAQ | 674500 | 0.80 | 0.62 |
| | AMZN | Amazon Corp. | NASDAQ | 1215700 | SBUX | Starbucks Corp. | NASDAQ | 1287200 | 0.80 | 0.45 |
| | SPLS | Staples Inc. | NASDAQ | 674500 | SBUX | Starbucks Corp. | NASDAQ | 1287200 | 0.81 | 0.69 |
| Consumer Staples | KO | Coca Cola Co. | NYSE | 1152559200 | SLE | Sara Lee Corp. | NYSE | 542934700 | 0.42 | 0.90 |
| | SLE | Sara Lee Corp. | NYSE | 542934700 | WMT | Wal-Mart Stores | NYSE | 1992433400 | 0.43 | 1.02 |
| | CVS | CVS Caremark Corp. | NYSE | 1454768200 | KFT | Kraft Foods Inc-A | NYSE | 1438368700 | 0.44 | 0.98 |
| | KFT | Kraft Foods Inc-A | NYSE | 1438368700 | SLE | Sara Lee Corp. | NYSE | 542934700 | 0.48 | 0.87 |
| | COST | Costco Co. | NASDAQ | 237500 | WFMI | Whole Foods Market | NASDAQ | 211400 | 0.59 | 1.09 |
| Energy | EP | El Paso Corp. | NYSE | 697103700 | RDC | Rowan Cos. | NYSE | 369346200 | 0.39 | 0.91 |
| | CVX | Chevron Corp. | NYSE | 1271849400 | XOM | Exxon Mobil Corp. | NYSE | 2798325600 | 0.41 | 1.00 |
| | XOM | Exxon Mobil Corp. | NYSE | 2798325600 | HAL | Halliburton Co. | NYSE | 1701703200 | 0.45 | 1.18 |
| | CVX | Chevron Corp. | NYSE | 1271849400 | EP | El Paso Corp. | NYSE | 697103700 | 0.45 | 0.88 |
| | COP | ConocoPhillips | NYSE | 1382115600 | EP | El Paso Corp. | NYSE | 697103700 | 0.46 | 0.94 |
| Financials | CINF | Cincinnati Financial | NASDAQ | 55900 | TROW | T. Rowe Price Group | NASDAQ | 216400 | 0.60 | 1.74 |
| | FITB | Fifth Third Bancorp | NASDAQ | 218000 | HBAN | Huntington Bancshares | NASDAQ | 71600 | 0.62 | 2.01 |
| | FITB | Fifth Third Bancorp | NASDAQ | 218000 | TROW | T. Rowe Price Group | NASDAQ | 216400 | 0.63 | 1.34 |
| | SCHW | Charles Schwab | NASDAQ | 1445500 | ETFC | E*Trade Financial Corp. | NASDAQ | 1391800 | 0.67 | 0.44 |
| | TROW | T. Rowe Price Group | NASDAQ | 216400 | ZION | Zions Bancorp | NASDAQ | 48500 | 0.70 | 1.40 |
| Health Care | BIIB | BIOGEN IDEC Inc. | NASDAQ | 381800 | PDCO | Patterson Cos. Inc. | NASDAQ | 106400 | 0.60 | 1.15 |
| | BIIB | BIOGEN IDEC Inc. | NASDAQ | 381800 | GENZ | Genzyme Corp. | NASDAQ | 242900 | 0.62 | 0.85 |
| | GENZ | Genzyme Corp. | NASDAQ | 242900 | GILD | Gilead Sciences | NASDAQ | 275100 | 0.64 | 0.91 |
| | CELG | Celgene Corp. | NASDAQ | 619200 | ESRX | Express Scripts | NASDAQ | 998400 | 0.65 | 0.47 |
| | BSX | Boston Scientific | NYSE | 1205569800 | THC | Teneo Healthcare Corp. | NYSE | 805228900 | 0.68 | 0.60 |
| Industrials | GE | General Electric | NYSE | 4303823300 | LUV | Southwest Airlines | NYSE | 862775700 | 0.53 | 0.73 |
| | CTAS | Cintas Corporation | NASDAQ | 48500 | EXPD | Expeditor Int'l | NASDAQ | 215600 | 0.54 | 1.48 |
| | EXPD | Expeditors Int'l | NASDAQ | 215600 | MNST | Monster Worldwide | NASDAQ | 196700 | 0.55 | 1.20 |
| | CHRW | C.H. Robinson Worldwide | NASDAQ | 112400 | EXPD | Expeditors Int'l | NASDAQ | 215600 | 0.56 | 1.46 |
| | MNST | Monster Worldwide | NASDAQ | 196700 | PCAR | PACCAR Inc. | NASDAQ | 164400 | 0.56 | 1.23 |
| Information Technology | DELL | Dell Inc. | NASDAQ | 909400 | ORCL | Oracle Corp. | NASDAQ | 1731900 | 0.79 | 0.47 |
| | CSCO | Cisco Systems | NASDAQ | 4886800 | DELL | Dell Inc. | NASDAQ | 909400 | 0.79 | 0.41 |
| | INTC | Intel Corp. | NASDAQ | 5529100 | ORCL | Oracle Corp. | NASDAQ | 1731900 | 0.82 | 0.29 |
| | CSCO | Cisco Systems | NASDAQ | 4886800 | ORCL | Oracle Corp. | NASDAQ | 1731900 | 0.83 | 0.34 |
| | AAPL | Apple Inc. | NASDAQ | 4627500 | INTC | Intel Corp. | NASDAQ | 5529100 | 0.84 | 0.19 |
| Materials | DD | Du Pont (E.I.) | NYSE | 716944300 | FCX | Freeport-McMoran Cp & Gld | NYSE | 1058215000 | 0.36 | 1.02 |
| | FCX | Freeport-McMoran Cp & Gld | NYSE | 1058215000 | MON | Monsanto Co. | NYSE | 479605700 | 0.36 | 0.85 |
| | APD | Air Products & Chemicals | NYSE | 187953500 | BLL | Ball Corp. | NYSE | 119560600 | 0.36 | 0.88 |
| | ECL | Ecolab Inc. | NYSE | 163404500 | NEM | Newmont Mining Corp. (Hldg. Co.) | NYSE | 958900000 | 0.37 | 0.93 |
| | FCX | Freeport-McMoran Cp & Gld | NYSE | 1058215000 | NEM | Newmont Mining Corp. (Hldg. Co.) | NYSE | 958900000 | 0.43 | 1.06 |
| Telecommunication Services | T | AT&T Inc. | NYSE | 2663617200 | Q | Qwest Communications Int | NYSE | 1623807700 | 0.45 | 1.01 |
| | S | Sprint Nextel Corp. | NYSE | 2044634000 | VZ | Verizon Communications | NYSE | 1472335800 | 0.49 | 0.87 |
| | Q | Qwest Communications Int | NYSE | 1623807700 | VZ | Verizon Communications | NYSE | 1472335800 | 0.49 | 0.86 |
| | T | AT&T Inc. | NYSE | 2663617200 | VZ | Verizon Communications | NYSE | 1472335800 | 0.50 | 1.05 |
| | Q | Qwest Communications Int | NYSE | 1623807700 | S | Sprint Nextel Corp. | NYSE | 2044634000 | 0.52 | 0.84 |
| Utilities | DUK | Duke Energy | NYSE | 902519200 | TE | TECO Energy | NYSE | 177983100 | 0.35 | 0.87 |
| | D | Dominion Resources | NYSE | 321656100 | XEL | Xcel Energy Inc | NYSE | 337262500 | 0.36 | 0.92 |
| | CMS | CMS Energy | NYSE | 264225200 | DUK | Duke Energy | NYSE | 902519200 | 0.39 | 0.81 |
| | CMS | CMS Energy | NYSE | 264225200 | DYN | Dynegy Inc. | NYSE | 702035600 | 0.48 | 0.79 |
| | DUK | Duke Energy | NYSE | 902519200 | DYN | Dynegy Inc. | NYSE | 702035600 | 0.54 | 0.74 |

References

- [1] T. W. Epps, Comovements in stock prices in the very short run, *Journal of the American Statistical Association* 74 (366) (1979) 291–298.
URL <http://www.jstor.org/stable/2286325>
- [2] G. Bonanno, F. Lillo, R. N. Mantegna, High-frequency cross-correlation in a set of stocks, *Quantitative Finance* 1 (1) (2001) 96–104.
- [3] J. Kwapien, S. Drozdz, J. Speth, Time scales involved in emergent market coherence, *Physica A* 337 (1-2) (2004) 231 – 242.
doi:10.1016/j.physa.2004.01.050.
URL <http://www.sciencedirect.com/science/article/B6TVG-4BRJY2K-1/2/d141755cb92c5d1b33cead1bb57b0bed>
- [4] M. Tumminello, T. D. Matteo, T. Aste, R. Mantegna, Correlation based networks of equity returns sampled at different time horizons, *The European Physical Journal B* 55 (2) (2007) 209–217.
doi:10.1140/epjb/e2006-00414-4.
URL <http://dx.doi.org/10.1140/epjb/e2006-00414-4>
- [5] A. A. Zebedee, M. Kasch-Haroutounian, A closer look at co-movements among stock returns, *Journal of Economics and Business* 61 (4) (2009) 279 – 294.
doi:10.1016/j.jeconbus.2008.11.001.
URL <http://www.sciencedirect.com/science/article/B6V7T-4TX794N-1/2/26edc1f881adc4aad35511cb584cb8d2>
- [6] M. Lundin, M. Dacorogna, U. A. Müller, *Financial Markets Tick By Tick*, Wiley, 1998, Ch. Correlation of High Frequency Financial Time Series.
- [7] J. Muthuswamy, S. Sarkar, A. Low, E. Terry, Time variation in the correlation structure of exchange rates: high-frequency analyses, *Journal of Futures Markets* 21 (2) (2001) 127–144.
- [8] R. Schäfer, M. Sjölin, A. Sundin, M. Wolanski, T. Guhr, Credit risk—a structural model with jumps and correlations, *Physica A: Statistical Mechanics and its Applications* 383 (2) (2007) 533 – 569.
doi:10.1016/j.physa.2007.04.053.
URL <http://www.sciencedirect.com/science/article/B6TVG-4NKJ0FV-4/2/b5717970f8ab98dedb2034f3c64ab7af>
- [9] R. Schäfer, N. Nilsson, T. Guhr, Power mapping with dynamical adjustment for improved portfolio optimization, *Quantitative Finance* doi:10.1080/14697680902748498.
URL <http://www.informaworld.com/10.1080/14697680902748498>
- [10] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertész, A. Kanto, Dynamics of market correlations: Taxonomy and portfolio analysis, *Physical Review E* 68 (5) (2003) 056110.
doi:10.1103/PhysRevE.68.056110.
- [11] T. Hayashi, N. Yoshida, On covariance estimation of non-synchronously observed diffusion processes, *Bernoulli* 11 (2) (2005) 359–379.
- [12] V. Voev, A. Lunde, Integrated Covariance Estimation using High-frequency Data in the Presence of Noise, *Journal of Financial Econometrics* 5 (1) (2007) 68–104. arXiv:<http://jfec.oxfordjournals.org/cgi/reprint/5/1/68.pdf>, doi:10.1093/jjfinec/nbl011.
URL <http://jfec.oxfordjournals.org/cgi/content/abstract/5/1/68>
- [13] J. E. Griffin, R. C. Oomen, Covariance measurement in the presence of non-synchronous trading and market microstructure noise, *Journal of Econometrics* In Press, Corrected Proof.
- [14] R. A. Edelson, J. H. Krolik, The discrete correlation function - A new method for analyzing unevenly sampled variability data, *Astrophysical Journal* 333 (1988) 646–659.
doi:10.1086/166773.
- [15] F. Corsi, F. Audrino, Realized Correlation Tick-By-Tick, *SSRN eLibrary*.
URL <http://ssrn.com/paper=957997>
- [16] L. Zhang, Estimating covariation: Epps effect, microstructure noise, *Journal of Econometrics* In Press.
doi:10.1016/j.jeconom.2010.03.012.
- [17] O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, N. Shephard, Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equities, *Economics Papers* 2008-W10, Economics Group, Nuffield College, University of Oxford (2008).
URL <http://ideas.repec.org/p/nuf/econwp/0810.html>
- [18] R. Renò, A closer look at the epps-effect, *International Journal of Theoretical and Applied Finance* 6 (1) (2003) 87–102.
- [19] B. Tóth, J. Kertész, The epps effect revisited, *Quantitative Finance* 9 (7) (2009) 793–802.
doi:10.1080/14697680802595668.
- [20] O. E. Barndorff-Nielsen, N. Shephard, Econometric analysis of realised covariation: High frequency covariance, regression and correlation in financial economics, *Economics Papers* 2002-W13, Economics Group, Nuffield College, University of Oxford (Nov. 2001).
URL <http://ideas.repec.org/p/nuf/econwp/0213.html>
- [21] W. F. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk, *The Journal of Finance* 19 (3) (1964) 425–442.
URL <http://www.jstor.org/stable/2977928>
- [22] J. D. Noh, Model for correlations in stock markets, *Physical Review E* 61 (5) (2000) 5981–5982.
doi:10.1103/PhysRevE.61.5981.
- [23] P. Gopikrishnan, V. Plerou, L. A. Nunes Amaral, M. Meyer, H. E. Stanley, Scaling of the distribution of fluctuations of financial market indices, *Physical Review E* 60 (5) (1999) 5305–5316.
doi:10.1103/PhysRevE.60.5305.
- [24] J. Voit, From brownian motion to operational risk: Statistical physics and financial markets, *Physica A: Statistical Mechanics and its Applications* 321 (1-2) (2003) 286 – 299.
doi:10.1016/S0378-4371(02)01783-1.
URL <http://www.sciencedirect.com/science/article/B6TVG-47CJ9HW-D/2/f8f724aa49a6da1bdbd15ed143ef7e42>
- [25] R. F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation, *Econometrica* 50 (4) (1982) 987–1007.
URL <http://www.jstor.org/stable/1912773>
- [26] R. F. Engle, Garch 101: The use of arch/garch models in applied econometrics, *Journal of Economic Perspectives* 15 (4) (2001) 157–168.
URL <http://ideas.repec.org/a/aea/jecper/v15y2001i4p157-168.html>
- [27] New York Stock Exchange, *Monthly TAQ DVD*, License #400570 (2007).